DOCUMENT RESUME

ED 087 459 IR 000 203

AUTHOR Harvey, Glenn B.
TITLE Rounding Second Base With Data Base: DePaul University's Experience.
INSTITUTION De Paul Univ., Chicago, Ill.
PUB DATE Apr 73
NOTE 16p.; Paper presented at the Association for Educational Data Systems Annual Convention (New Orleans, Louisiana, April 16 through 19, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Computer Oriented Programs; *Computer Programs; *Data Processing; Higher Education; *Management; Program Descriptions
IDENTIFIERS AEDS; Association for Educational Data Systems; *Data Base; De Paul University; Generalized Retrieval System

ABSTRACT

A description of Data Base, the computerized data processing system used by De Paul University, is provided. Some background information on the school's data processing experiences and its decision to adopt Data Base are first presented, followed by a brief review of the Data Base software selected for use by the University. Fuller details are then given about the data management system, including information about its job stream, construction standards, dictionary of elements, file relationships, and hardware. Lastly, a description of Data Base's Generalized Retrieval System is offered. The system's main components and characteristics are outlined and an illustration of its operation is provided. (PB)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 087459

# ROUNDING SECOND BASE WITH DATA BASE
## DePAUL UNIVERSITY'S EXPERIENCE

Glenn B. Harvey

DePaul University (Illinois)

## INTRODUCTION - DePAUL'S EARLY COMPUTER ENVIRONMENT

DePaul University has always followed a tight budget philosophy toward computers and data processing. This is altogether consistent with the University's goal of providing a cost effective program in higher education.

DePaul has approximately 10,000 students, over 525 faculty and over 40,000 graduate alumni. An additional 250,000 people have taken educational programs of some duration at the University.

The 1972-73 operating budget is approximately $16.5 million, nearly triple the $5.5 million budget of 1963-64.

Strict control on the funding of computers paid financial dividends but left a widening instructional gap for DePaul graduates who learned little about the electronic marvel. On the administrative side, second generation IBM 1401 equipment, software and concepts were relied upon, even into the 1970's when surrounding institutions were processing from third to fourth generation philosophies. Because operations staff and programming resources were also held to a minimum, this effectively curtailed DePaul's ability to respond to changing needs.

Pressures from both internal and external sources became stronger to expand the computer resources. By late 1971, this resulted in the acquisition of new high-level experienced in the application of computers to both administrative and instructional systems and dedicated to the proposition that computer systems should service users in unique new ways.

## DECISION TO GO DATA BASE

To lend continuity and control to Systems Planning and Development efforts, a Master Plan for University-wide Computer Services was written and presented to the Executive Committee for approval. This plan outlined the alternatives for computer services available to DePaul, their associated costs and benefits.

Included were recommendations for hardware, software, staff and an implementation schedule for application systems. Objectives to be achieved by new systems required that they be:

---

GLENN B. HARVEY is Manager of Administrative Systems Development for Education, Honeywell, Inc., Minneapolis. He holds a B.S. and M.S. (Business Administration) from the University of Illinois and a Certificate of Data Processing.



CE 000 803

* Responsive - To changes in user needs regardless of timing of of source of change.
* Flexible - Able to provide for new files, new data in old files, new methods of processing old data, entire new systems, or whatever need arises.
* Simple - Easily understood by users of data, as well as the computer technician.
* Self-documenting - System should provide relatively complete documentation of itself as a by-product of its own operation.
* Parameter-driven - Most functions to be handled by entry of control cards or tables with tailored programming held to a bare minimum.
* User-controlled - User should have high degree of control over his own data files, their content, maintenance and usage.
* Easily-converted - Little difficulty converting existing 1401 files from old to new systems.
* Economical - All of the above goals should be possible on relatively low-cost computer equipment with small memory and few input/output units. Staff expenditures should likewise be minimal yet able to accommodate system demands and changes.

It was recommended and approved that the new machine system* be constructed using Data Base concepts. That is, files within the "data base" would be logically inter-related in a network fashion. This had the obvious advantage of permitting the development of application systems that could easily retrieve new "relationships" of data, i.e., new information for new uses, rather than just the data themselves. Traditional file-oriented machine systems are programmed to contain specific data relationships that satisfy needs at a point in time, but are time consuming and costly to reprogram when changes occur in the needs for information.

## SELECTION OF DATA BASE TOOLS

Several decisions were also made in regard to existing computer software that could aid DePaul's development of a Data Base system. Much time was devoted to seeking out and evaluating "packages" that purported to cure an organization's data processing problems in information retrieval, report writing, total information systems, management information systems, data base management, data management or file management systems. Some significant discoveries were made.

This paper does not recommend nor condemn any given "package." We looked into many software aids that were rejected, for DePaul's use, for one reason or another. Included in this study were products and names such as ASSIST, BUG, Burrough's DISK FORTE, DATA ANALYZER, IBM's IMS & EPIC, MARK IV, Meta-COBOL, PRISM, RSVP, SCORE III, Stanford's "OASIS," SERIES, WORK TEN and others. Many of these products have excellent features, but still did not meet our criteria. This sometimes

* Hardware selected was an interim System 360/30, 96K memory, 4 tape/3 disk, computer to be followed by an IBM System 370/135, with the same memory input/output units, when available in January 1973. The 3 disks were originally 2319 replaced by double-density disks in August, 1972, thus making approximately 175 million bytes of storage available on-line. The interim hardware was installed March, 1972.

meant they cost too much. While a separate paper could be written on such an evaluation, several observations should be made here:

* Price is often negotiable, or subject to change.
* No one really knows what software is worth.
* A single user's recommendation for a package is never enough... sometimes several isn't enough.
* Good, cheap packages exist that are virtually undiscovered and unknown.

While investigating one such product (a data base management system) called TOTAL, we learned of its use at Amherst College in Amherst, Mass. Amherst had developed a sophisticated generalized maintenance system using TOTAL as the "super-access" method for organizing its disk files. Their approach to data processing problems appeared to us more advanced, responsive, flexible, etc. than any other we found. That system (or systems), used with the TOTAL package, seemed to be DePaul's answer to new system development. The Amherst system became the foundation for the DePaul system herein described. Only because Amherst provided their programs, documentation, staff time and patience while we were learning their operation, has the DePaul capability been made possible.

We believe that this has been one of the best examples of sharing computer systems between institutions that one can find. All of DePaul's enhancements to the system have, in return, been given to Amherst in token payment. Cost of the base system actually included (1) acquisition of the TOTAL package from its vendor, Cincom Systems, Inc. of Cincinnati, (2) staff time and expenses to visit Amherst on several occasions, (3) staff and machine time to study, analyze and modify the system from RCA to IBM computers and slightly different versions of COBOL, and (4) staff and machine time to enhance, in some cases rewrite, original programs to do more using different techniques.

Other software products DePaul acquired at this time included a COBOL Program Generator called SNAP (Simplified Notation for Application Programmers) that permits us to produce accurate, low-core, cross-reference COBOL D programs in a quick and easy manner. All of the possible features and benefits of this excellent, low-cost (under $3000) package have yet to be used by DePaul.

For our entry into the world of Teleprocessing, we selected a package called CRT-Interface, marketed by Westinghouse Tele-Computer Systems of Pittsburgh. This package requires minimal memory for operation (32K partition), permits COBOL programming of TP jobs and cost DePaul $900. To facilitate our data base back-up, we also purchased the Westinghouse Dump/Restore disk-to-tape program, for $800, the best program of its kind.

## DATA BASE SYSTEM STANDARDS

The DePaul System is really a Data Management System. It is composed of four basic parts:

* Generalized Maintenance System (GMS)
* Generalized Retrieval (and Reporting) System (GRS)
* Generalized Utility System (GUTS)

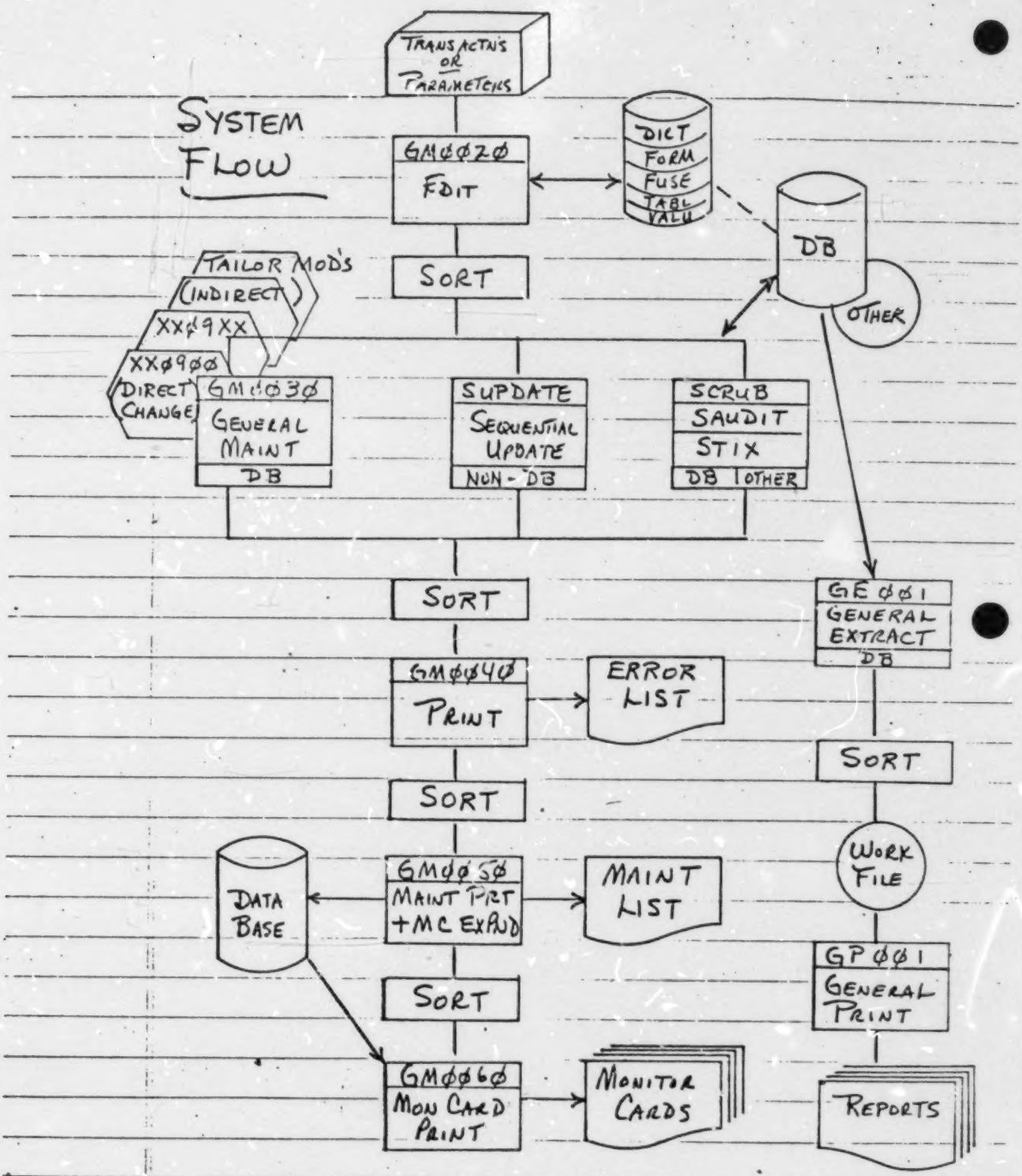

SYSTEM FLOW
TRANSACTNS OR PARAMETERS
GM ØØ2Ø EDIT
DICT
FORM
FUSE
TABL
VALU
DB
OTHER
SORT
TAILOR MOD's
(INDIRECT)
XXØ9XX
XXØ9ØØ
DIRECT CHANGE
GM ØØ3Ø GENERAL MAINT DB
SUPDATE SEQUENTIAL UPDATE NON-DB
SCRUB SAUDIT STIX DB OTHER
SORT
GE ØØ1 GENERAL EXTRACT DB
GM ØØ4Ø PRINT
ERROR LIST
SORT
SORT
WORK FILE
DATA BASE
GM ØØ5Ø MAINT PRT +MC EXPND
MAINT LIST
GP ØØ1 GENERAL PRINT
SORT
GM ØØ6Ø MON CARD PRINT
MONITOR CARDS
REPORTS

FIGURE 1

The fourth part is being added right now and permits TP inquiry/update into the Data Base -- to be called Generalized Terminal System (GTS). This paper treats only the second system, GRS, our method of retrieving information from the data base files.

GMS Job Stream

A single job stream GMJ001 exists for the updating, ala maintenance, of any and all files in the Data Base. All files can be active in the same run, that is, incoming transactions can be targeted for any Data Base file. It is expected that only one GMS run will be required each day to maintain all such files. The System Flow in Figure 1 shows input transactions being converted to disk via the IBM card-to-disk utility. Program GM0020 edits each transaction thoroughly, flags those with errors and releases them all to the update program GM0030. GM0030 is a calling program that calls the generalized module XX0900 to process all direct changes to the appropriate Data Base files.

A direct change transaction uses one of the transaction codes shown in Figure 2. In effect, such transactions affect no file or item of data except that which is specified in the transaction itself. An example would be the recording of a student being registered in a specific course.

GM0030 can also call numerous other update modules for the purpose of making indirect changes. Tailored programs, such modules are triggered by the type of transaction entered into the job stream. Thus, in any given run, many different types of files, activity, and indirect changes can be involved. An example of an indirect change could be the setting of an Enrollment Status Indicator, in a different file, when a student enrolls for any course.

GM0040 prints all error, unposted transactions that entered a given run, while GM0050 prints the posted conditions. In addition, GM0050 analyzes the users of data posted and generates, to multiple users if necessary, monitor reports of records that have changed.

Parameter-driven GRS

At any time, and at regularly scheduled reports, the GRS capability permits extraction of records from one or more Data Base files with or without sorting, for the printing of reports (lists, labels, accounting reports, statistics, etc.) all by parameter entry. It is expected that at least 80% of the computer center's output will be handled by the GRS -- if the Amherst installations is a good example.

Data Base Standards of Construction

Constraints on construction of the DePaul Data Base were established in order to use the GMS to maintain all files with but one update module. First, some definitions are in order.

* Physical File = Grouping of similiar records within the Data Base. TOTAL permits two types, master or single-entry and variable or variable-entry. Each is identified to TOTAL by a four-letter name, such as PEOP (for a PEOPLE file of names, etc.).

* File Record = Logically unique data grouping for a given primary control (like Soc. Sec. #) in the case of a master file. A variable file record needs both a primary control segment and a secondary control (part of SEG1) to establish its uniqueness -- and thus GMS can find the record when trying to update it.

* Segment - Grouping of data elements within a file record. CTRL segment contains only the primary control element. A data segment, i.e. SEG1, can contain any number of data elements.

* Data Element = Usually the smallest data component of a segment. System permits multiple definition of elements.

While the GMS programs could be modified to accommodate almost any type of Data Base system, the setting of standards within our own has greatly facilitated the development of new methods and techniques for dealing with the data contained therein.

Data Base Dictionary of Elements

Several files in the Data Base are used solely for control of the system. Perhaps the most important is the DICT or Dictionary of Data Elements File. Every element is described in DICT, its associated file, segment, beginning position, size, class, etc. Each element has a unique number and is structured in a special way as seen in Figure 2. The element in the DePaul system is related to a physical file by its prefix.

In the final analysis, the DICT even describes itself, as seen in Figure 4. Here, it can be seen that 22 elements comprise one DICT record, 21 of which are in SEG1 and 1 in CTRL segment. A 23rd element really exists in DICTSEG1 positions 1 - 6 known as "last transaction date" which is automatically supplied and maintained by the GMS. Because the DICT can describe itself, the GMS can be used to update the DICT records of any element. In other words, changes to the structure of the system are as simple as changes to the data.

Data Base File Relationships

Another system control file FILS describes the characteristics of each physical file and its relationship to others in the Data Base; see Figure 3 for a sample Data Base Schematic and Figure 4 for the associated content of the FILS file. An interesting fact is that FILS can exist under any data base access method (not necessarily TOTAL). File relationships are also carried in the TOTAL core-resident descriptor module, but that is frequently proprietary, changeable or the vendor may even be changed. Like other data base files, FILS can be modified through GMS. FILS and DICT along with the attendant standards of data base construct, terminology and file relationships, form the heart of the DePaul system. Other system control files are explained later.

Hardware Considerations

The system is designed to occupy no more than 48K of memory for any program module. For instance, the update phase GM0030 can include in memory at one time:

## STANDARD DATA BASE TERMINOLOGY

* System No. = AA 2-digit alpha
  * System Logical File = Usually 1 mast file with some linked Var Files
  * Such as RG = Registrar
    AD = Admissions
* File No. = AA (Standard Abbreviation) or AAAA ('Total' Name)
  * Physical File = Master or Variable
  * Such as PE = PEOP
    CR = COUR
* Data Element No. = AA (File Abbreviation) NNN (Element #)
  * Identified by Physical File
  * AAØØ1 required for Master or Variable Files and = Control # (up to 30)
  * Such as PEØØ1 = PEOP SS#
    PEØ1Ø = PEOP Last Name
  * Can be redefined in DICT to designate combinations of other elements or the unique processing of the given element.

## STANDARDS FOR DATA BASE CONSTRUCTION

1. Each Physical File Record =
   * Control Segment (up to 30 characters)
   * Data Segment(s) (up to 240 characters)
     * Master File can have 1 to 9 segments
     * Variable File can have 1 segment
     * Segment Name is "SEGn" (e.g. SEG1)
2. Master File
   * # of links to Variable Files = 1 to 2Ø
   * First 6 digits SEG1 = Last Trans Date
3. Variable File
   * All updating done via a single Base Master File
   * Must have Record Code (for TOTAL)
   * Has Record Type first 2 digits SEG1
   * Can have additional secondary control following Record Type

## TRANSACTION CODES

AA = Add Master Record

DD = Delete Master Record and any associated Variable Files

CR = Change element in Master or Variable Record by Replacement

-or-

Create new Variable Record if none exists and insert element value from this transaction.

CE = Change element in Master or Variable Record by Erasure (set to blanks or zeroes as appropriate)

-and-

Completely eliminate a Variable Record if this was its last data element remaining.

CA = Change element by numeric addition or subtraction.

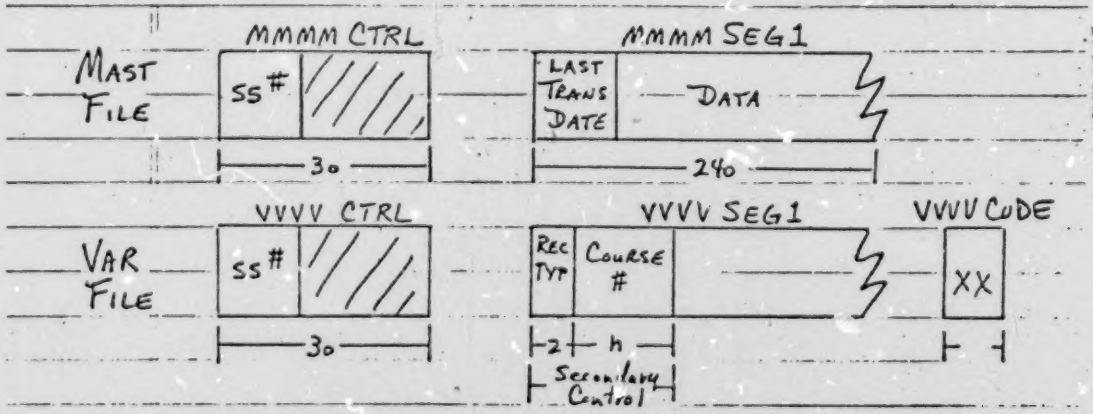
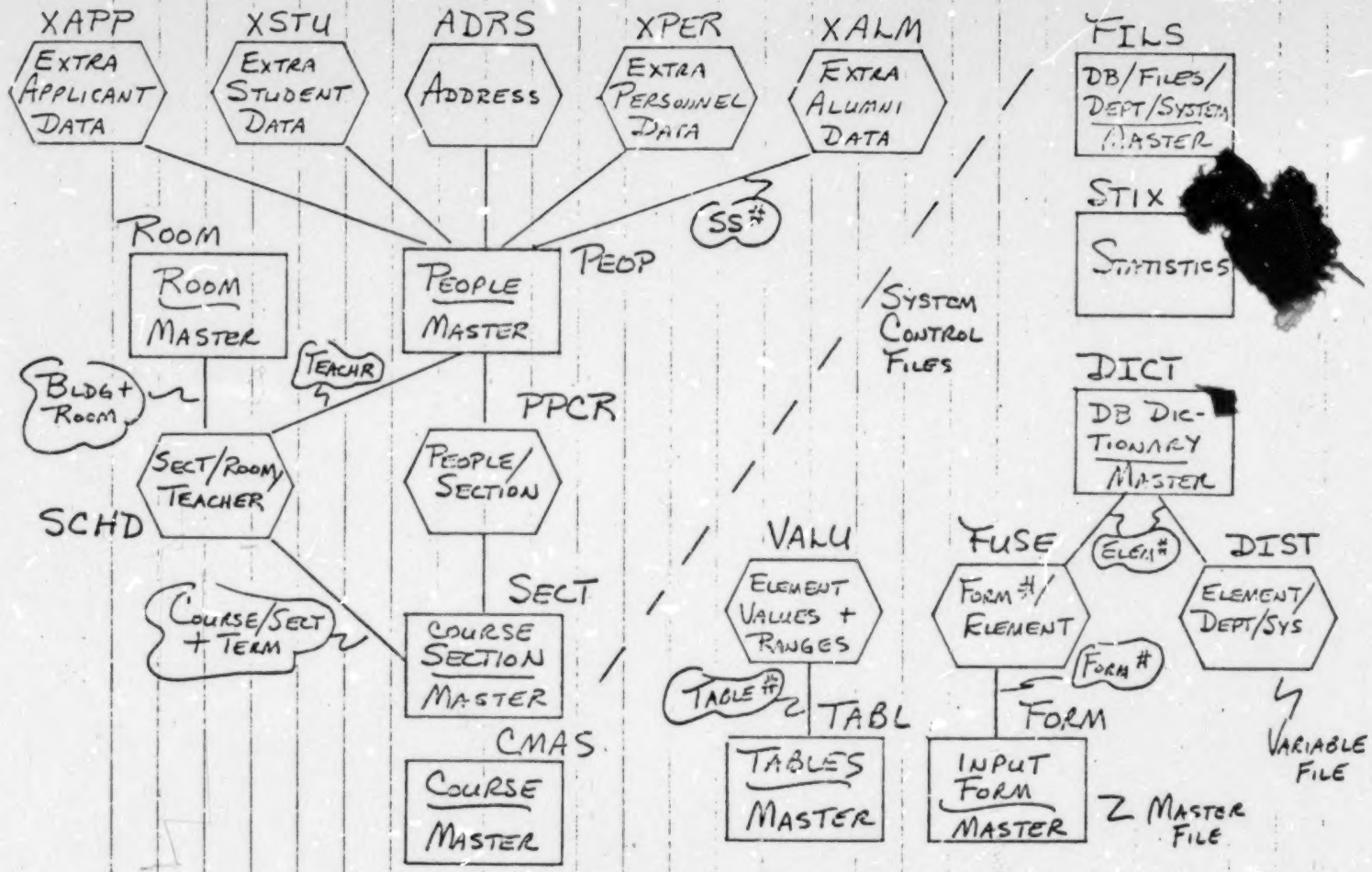



FIGURE 2

557




DEPAUL DATA BASE

FIGURE 3

INITIAL LOAD OF DATA BASE FILS DATA ELEMENTS 09/01/72

| NO. | M/V | SGS | LKS | UPDT LK | LK | LK01 LK02....ETC. |
|---|---|---|---|---|---|---|
| ..... | . | . | .. | .... | .. | .... .... .... |
| FILS | M | 1 | 00 | | 0 | |
| DICT | M | 1 | 00 | | 0 | |
| PEOP | M | 2 | 02 | | 0 | PPCR ADRS1 |
| COUR | M | 1 | 01 | | 0 | PPCR |
| PPCR | V | 1 | 0 | PEOP | 01 | |
| ADRS | V | 1 | 0 | PEOP | 02 | |
| MF01 | | 0 | 04 | | 0 | FILS DICT PEOP COUR |
| VF01 | | 0 | 02 | | 0 | PPCR ADRS |
| PE0114 | | B | 2 | 007020 | 027003 | |
| PC0114 | | B | 2 | 007020 | 027003 | |
| AD0114 | | B | 2 | 007020 | 027003 | |
| PE0115 | | A | 3 | 007012 | 030004 | 027003 |
| CR0114 | | A | 1 | 007015 | | |

INITIAL LOAD OF DATA BASE DICTIONARY DATA ELEMENTS 09/01/72

| NO. | M/V | RCOD | FILE/SEG | SIZE | LOC | CLS | DEC | CHG | NAME | DESCRIPTION | DEPT | R | SHIFT | UP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ..... | . | .. | ..... | ... | ... | . | . | ... | ..... | ..... | ... | . | ... | .. |
| DI001 | M | | DICTCTRL | 5 | 1 | X | | | DINUMB | DATA ELEMENT NUMBER | 114 | C | 000 | |
| DI020 | M | | DICTSEG1 | 1 | 7 | X | | ER | DITYPE | FILE TYPE | 114 | | 000 | |
| DI030 | M | | DICTSEG1 | 2 | 8 | X | | ER | DIRCOD | RECORD CODE - TOTAL | 114 | | 000 | |
| DI040 | M | | DICTSEG1 | 8 | 10 | X | | ER | DISEGM | SEGMENT NAME / FILE NAME | 114 | | 000 | |
| DI050 | M | | DICTSEG1 | 3 | 18 | N | 0 | ERA | DILENG | ELEMENT LENGTH %BYTES¤ | 114 | | 000 | |
| DI060 | M | | DICTSEG1 | 3 | 21 | N | 0 | ERA | DILOCA | ELEMENT LOCATION %LEFT¤ | 114 | | 000 | |
| DI070 | M | | DICTSEG1 | 1 | 24 | X | | ER | DICLAS | ELEMENT CLASS | 114 | | 000 | |
| DI080 | M | | DICTSEG1 | 1 | 25 | N | 0 | ERA | DIDECM | DECIMAL PRECISION | 114 | | 000 | |
| DI090 | M | | DICTSEG1 | 3 | 26 | X | | ER | DICHNG | VALID CHANGE TYPES | 114 | | 000 | |
| DI100 | M | | DICTSEG1 | 6 | 29 | X | | ER | DIMNEM | ELEMENT MNEMONIC | 114 | | 000 | |
| DI110 | M | | DICTSEG1 | 25 | 35 | X | | ER | DIDESC | ELEMENT DESCRIPTION | 114 | | 000 | |
| DI120 | M | | DICTSEG1 | 3 | 60 | N | 0 | ERA | DIDEPT | MAINTENANCE DEPARTMENT | 114 | | 000 | |
| DI130 | M | | DICTSEG1 | 1 | 63 | X | | ER | DIREST | CHANGE RESTRICTION CODE | 114 | | 000 | |
| DI140 | M | | DICTSEG1 | 3 | 64 | N | 0 | ERA | DISHIF | SHIFT LOC %AUTO¤ | 114 | | 000 | |
| DI150 | M | | DICTSEG1 | 3 | 67 | N | 0 | ERA | DIUPDT | UPDATE ROUTINE | 114 | | 000 | |
| DI160 | M | | DICTSEG1 | 2 | 70 | X | | ER | DITABL | VALIDATION TABLE NO. | 114 | | 000 | |
| DI170 | M | | DICTSEG1 | 2 | 72 | X | | ER | DIRTYP | RECORD TYPE - CODED REC. | 114 | | 000 | |
| DI180 | M | | DICTSEG1 | 3 | 74 | N | 0 | ERA | DISLOC | SECONDARY CONTROL LOC. LF | 114 | | 000 | |
| DI190 | M | | DICTSEG1 | 2 | 77 | N | 0 | ERA | DISLEN | SECONDARY CONTROL LENGTH | 114 | | 000 | |
| DI200 | M | | DICTSEG1 | 1 | 79 | X | | ER | DICCCR | MULTIPLE OCCURENCE CODE | 114 | | 000 | |
| DI210 | M | | DICTSEG1 | 1 | 80 | X | | ER | DIESSN | ESSENTIAL/NON-ESSENTIAL | 114 | | 000 | |
| DI220 | M | | DICTSEG1 | 1 | 81 | X | | ER | DIMONI | MONITOR CARD CODE | 114 | | 000 | |

INITIAL LOAD OF DATA BASE DICTIONARY DATA ELEMENTS 09/01/72

| NO. | M/V | RCOD | FILE/SEG | SIZE | LOC | CLS | DEC | CHG | NAME | DESCRIPTION | DEPT | R | SHIFT | UP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ..... | . | .. | ..... | ... | ... | . | . | ... | ..... | ..... | ... | . | ... | .. |
| PE001 | M | | PEOPCTRL | 9 | 1 | N | 0 | | PECTRL | PEOPLE FILE CONTROL NO. | 114 | C | 000 | 90 |
| PE010 | M | | PEOPSEG1 | 12 | 7 | X | | ER | PELNME | LAST NAME | 114 | | 000 | 90 |
| PE020 | M | | PEOPSEG1 | 8 | 19 | X | | ER | PEFNME | FIRST NAME | 114 | | 000 | 90 |
| PE030 | M | | PEOPSEG1 | 3 | 27 | X | | ER | PECLAS | CLASS | 114 | | 000 | 90 |
| PE040 | M | | PEOPSEG1 | 4 | 30 | X | | ER | PEMAJR | MAJOR | 115 | | 000 | 90 |
| PE050 | M | | PEOPSEG2 | 1 | 9 | N | 0 | ERA | PES01 | ALUMNI STATUS | 114 | | 000 | 90 |
| PE060 | M | | PEOPSEG2 | 1 | 10 | N | 0 | ERA | PES02 | ADMISSION STATUS | 114 | | 000 | 90 |
| PE070 | M | | PEOPSEG2 | 1 | 11 | N | 0 | ERA | PES03 | ACTIVE STUDENT STATUS | 114 | | 012 | 90 |
| PE080 | M | | PEOPSEG2 | 1 | 12 | N | 0 | ERA | PES03P | STUDENT STATUS-PREV | 114 | | 000 | 90 |
| PE090 | M | | PEOPSEG2 | 3 | 13 | X | | ER | PEGEOG | GEOGRAPHIC CODE | 114 | | 000 | 90 |
| PE100 | M | | PEOPSEG2 | 10 | 16 | X | | ER | PESTRT | STREET ADDR - LOCAL | 114 | | 000 | 90 |
| PE110 | M | | PEOPSEG2 | 8 | 26 | X | | ER | PECITY | CITY - LOCAL | 114 | | 000 | 90 |
| PE120 | M | | PEOPSEG2 | 2 | 34 | X | | ER | PESTAT | STATE - LOCAL | 114 | | 000 | 90 |
| PE130 | M | | PEOPSEG2 | 5 | 36 | X | | ER | PEZIP | ZIP CODE - LOCAL | 114 | | 000 | 90 |

FIGURE 4

* Root Module GM0030, GM1030, Macros
* TOTAL Data Base Description Module (Size varies)
* TOTAL Data Base Manager
* Transient COBOL modules like XX0900

A 65K IBM 360/30 easily accommodates this system at a very modest hardware cost. Programming language for all primary modules is IBM COBOL D. There are also 3 macros written in Assembler, each about 15 instructions. The system is thus highly hardware and software-independent. Operating system is DOS.

Our double-density ITEL disk drives are able to provide concurrent operations among (1) 1401 Emulation, (2) Data Base Batch Processing, (3) Data Base TeleProcessing, and (4) Spooling of I/O operations. See Figure 5 for hardware maps.

## DATA BASE INFORMATION RETRIEVAL - GRS

### System Components

The Generalized Retrieval (and Reporting) System (GRS) basically follows the functional sequence of processing steps in Figure 6 which shows the GRS Job Stream. Components of this include:

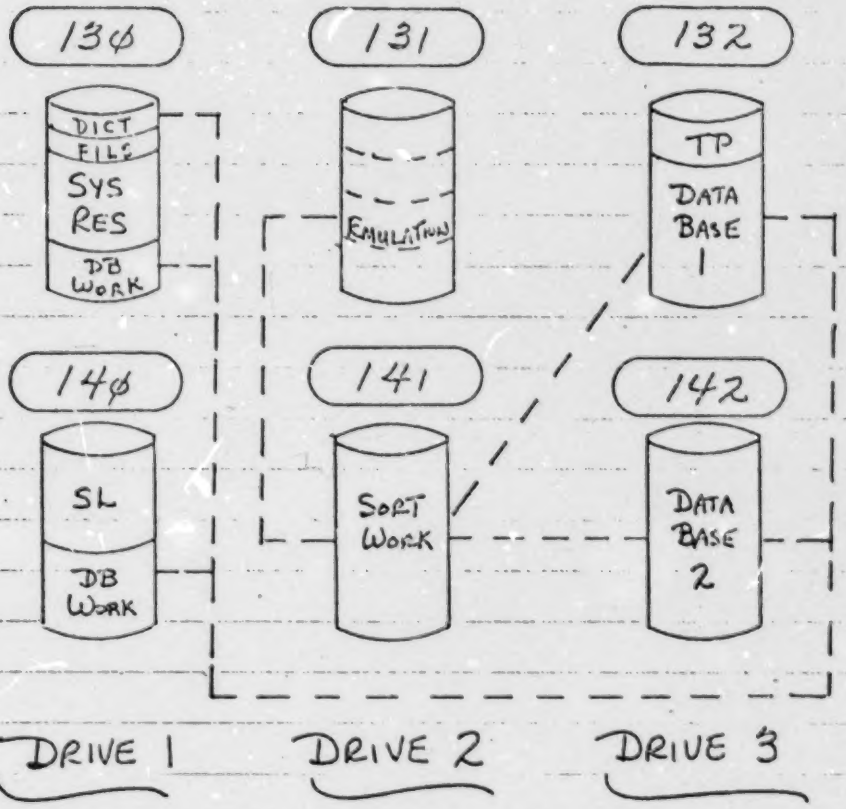
* Generalized Extract Program - Written in COBOL, this is the key to combining information from several data base files into one, or more, meaningful logical system outputs.
* Generalized Sort Program - Nothing more than the IBM Sort Utility, the first 60 characters of each record produced by the Extract are sequenced. This is the key to successful multiple-report processing because it places all records for a given report together, each with its own independent internal sequence (such as name within college)...according to whatever data fields the Extract placed in the 60-character sort key.
* Generalized Print Program - An infinite variety of printed outputs is possible from this, or collection of program modules known as General Print.

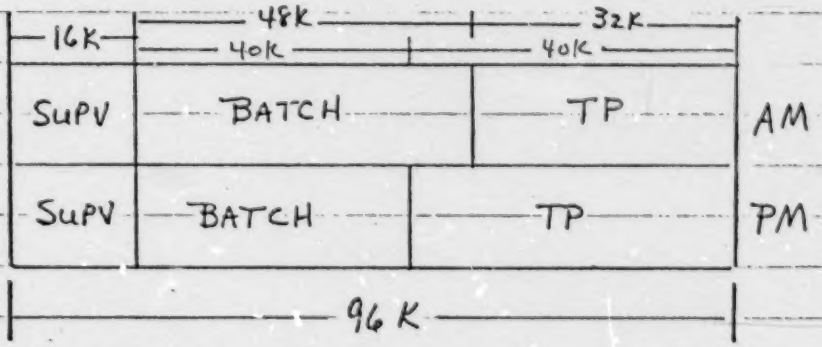
### System Characteristics

The GRS Job Stream can be used at any time to pass the desired data base files, extract data elements, sort and subsequently print them in any number of reports in any desired sequence and format. This, then, is the normal method for printing standard scheduled reports or other documents not associated with transactional activity to the data base. Special reports of an unscheduled nature can likewise be handled through this job stream.

The last phase of the GMS Job Stream is Monitor Record Generation, or the producing of specific records on which changes have taken place during the maintenance process. Monitoring is how a user knows what his computer data file looks like. It is as though the user had a "window" into his file which flicks the correct image of every record at him as he needs it.

## DISK LAYOUT




MEMORY LAYOUT

FIGURE 5

# GRS Job Stream

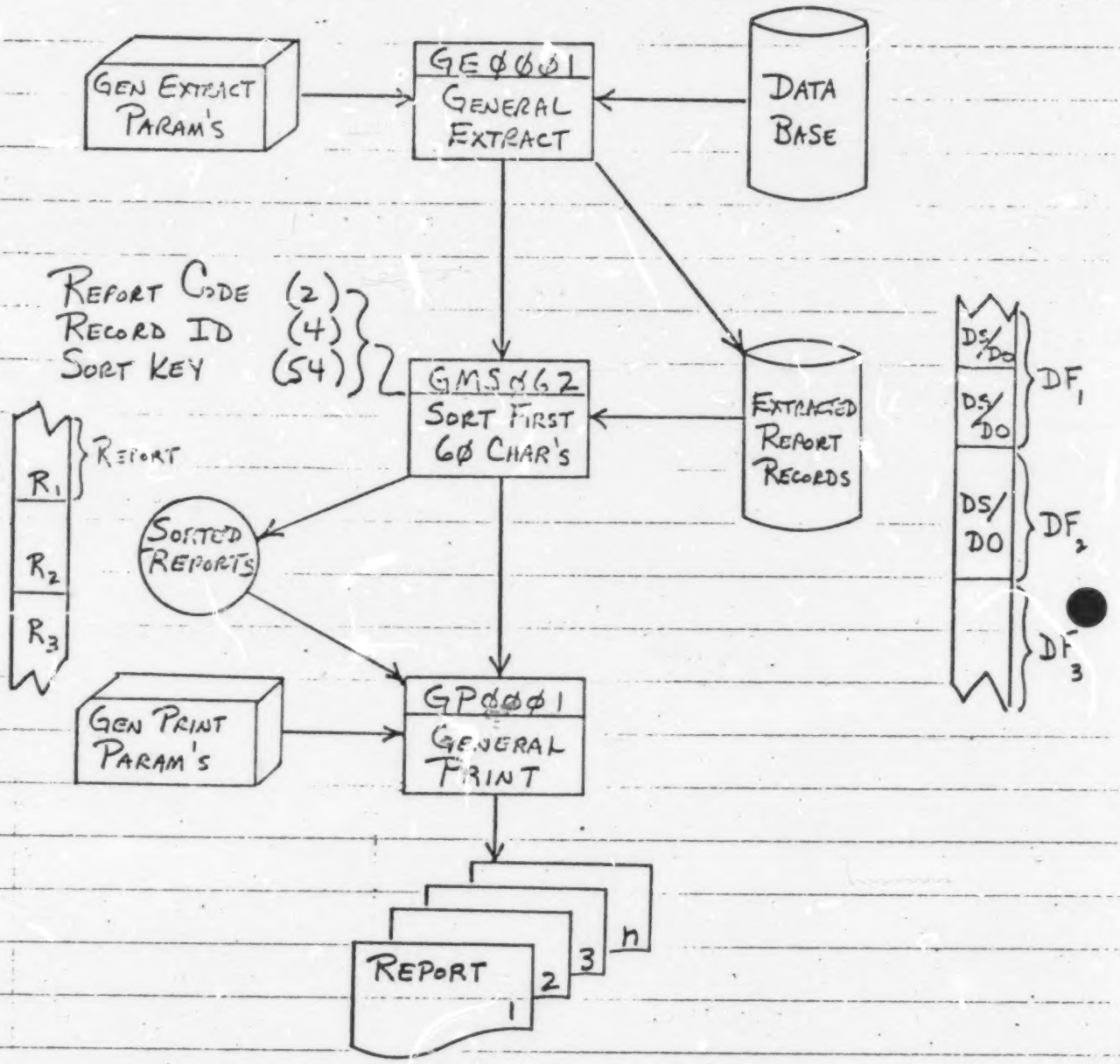



FIGURE 6

Both reports and monitor records are produced by the system components Generalized Extract, Sort and Print. It was mentioned earlier that all extraction and printing is accomplished by preparing and entering parameter cards (also known as control cards) to specify the details of what is to be processed and how it is to be sorted and printed. Some Extract parameters are:

* Fast and easy to prepare, but are oriented to the computer analyst, not the user.
* Optionally stored within the data base itself and retrieved by the GRS as needed to make data file extractions or prints.
* Operational by default. COBOL own-code program modules can be written to over-ride the generalized functions of the system, through own-code exits provided in GRS.

Efficient computer centers have known for years that, although difficult to program and implement, extracting many reports from a data file at one time, then sorting and printing them in one solid stream (more or less automatically) was inherently better than processing reports one at a time. The savings can be attributed to less job set-up time, better use of computer resources (like the printer), fewer opportunities for manual intervention and error, etc. GRS adds to multi-report capability the following features:

* Any number of reports, up to 40, can be processed in a single run of the job stream.
* Reports on the Generalized Print input tape can be printed or bypassed via a Report Selection Parameter...making possible any manner of re-starts, re-runs, or repeat prints.
* Reports to be printed are controlled against reports that were earlier extracted in an automatic sequence of console logging for operator and management use. A run summary is provided showing reports printed, bypassed, selected but missing, etc.

Simple extraction of all records from one or more data files does not constitute the retrieval of information. To be of any value, extensive record selection capabilities are necessary. Virtually any combination of selection criteria is permitted by both Generalized Extract and Print. In fact, records may even be selected by comparisons of several data fields within a given record, or by modification of certain data fields via table features explained below and then the result used for selection. In effect, selection is almost a miniature program written in parameter form and can be as simple or complex as one desires.

Either stock or custom pre-printed forms can be handled by GRS. A testline may be printed at the start of each report, as many times as necessary to align a special form...or optionally omitted, again, by report. Reports can be vertically and horizontally formatted in almost any fashion. Data from several records may be combined in a single print line or a single record may be spread vertically over any number of print lines...see sample Admissions Matriculation Card in Figure 7.

Complete editing of data fields, including designation of standard internal masks (or explicitly specified masks) and the emission of literals

is provided. Headings can include automatic dating and paging by report. In addition, table cards may be entered by report for the expansion or conversion of data fields prior to printing. Control totals for numeric fields and record counts based on the record selection logic (up to 63 record counts per report) can be generated for accounting reports and/or statistical summaries.

System Operations - A Multiple Report Example

The DePaul Data Base being used is shown in Figure 3 depicting its various files. The sample run is intended to generate 4 reports from 3 separate sweeps of the data base by General Extract, each report with its own selection criteria, format, sequence and content.

General Extract parameters are coded as DF (Define File), DO (Define Output) or DS (Define Selection) in char. 1-2. The DF group repeats for each sweep of the data base. DF cards identify the File Extraction Path the program will follow in order to retrieve data fields from files. Data fields thus retrieved will be used for testing purposes (according to the DS cards) or for output specifications (according to DO cards). DS and DO parameters must be related to each other by the identification in chars. 3-7 (e.g. TEST01). More than one DS/DO set can be entered for a group of DF cards.

Two reports, AA and AB, are being generated from the first data base pass by General Extract. Selection of females for the first report and males for the second is reflected in the respective DS cards. Note the data fields to be sorted are specified in the DO cards (such as in TEST01) where name element PE028 is being placed in position 0007 of the output record following the literal 'AA0101'.

Note that a data pass by the extract program involves only those data files specified in the DF cards, no others. Thus, extraction speed is limited to the time necessary to pass the disk files from beginning to end (sequential read). Multiple reports (DS/DO's) can, of course, be generated during that pass.

For cases where the specific records to be retrieved can be identified by their control number (like Soc Sec #), the extract program reads a "trigger" file of cards or tape containing the specific control numbers desired. A direct seek is then made to the records wanted and extraction is accomplished immediately. This is the technique used to generate Monitor Records from the GMS Job Stream each day.

After the sort, General Print accepts the sorted input tape along with a deck of print parameter cards, in report code sequence (e.g., AA,AB,XX,YY). These parameter cards are identified by char. 1 which can range from 1 to 6 as follows:

* Param 1 - Provides report code in chars. 2-3 along with input record length and blocking factor.
* Param 2 - Heading information. Can also contain standard line spacing and several other options for the report.
* Param 3 - Detail print line information including designators as to when line is to print and spacing before or after print. This card is also used to specify a Summary report line...not included in example.

DEPAUL PRINT GENERATOR EDIT/ERROR LIS MAR. 12, 1973 PAGE 0001

ADMISSIONS MONITOR MATRICULATION CARD

```
OFIXED
1N0050004
2000000000008
200100130A  X
3NON00010C   001003001    SOC SEC    (LINE 1)
3NON00020D   004-         EDITED     (LINE 1)
3NON00030C   005002004               (LINE 1)
3NON00040D   007-                    (LINE 1)
3NON00050C   008004006               (LINE 1)
3NON00060C   013030010    NAME       (LINE 1)
3NON00070C   045002044    BIRTHDAY   (LINE 1)
3NON00080D   047-         EDITED     (LINE 1)
3NON00090C   048002046               (LINE 1)
3NON00100D   050-                    (LINE 1)
3NON00110C   051002048               (LINE 1)
3NON00120C10 056002050    COLLEGE    (LINE 1)
3NON00200C   013020079    STREET     (LINE2)
3NON00210C   035014099    CITY       (LINE2)
3NON00220C   050002113    STATE      (LINE2)
3NON00230C12 053005115    ZIP        (LINE2)
3NON00250C   013030150    PARENT     (LINE 3)
3NON00260C   046003120M   PHONE      (LINE 3)
3NON00270C   050003123    EDITED     (LINE 3)
3NON00280D   053-                    (LINE 3)
3NON00290C12 054004126               (LINE 3)
3NON00300C   013028226    HI SCHOOL  (LINE 4)
3NON00310C   041002254    HS YEAR    (LINE 4)
3NON00320C   047001057    ENTRY TRM  (LINE 4)
3NON00330C12 049002058    ENTRY YR   (LINE 4)
```

SOC. SEC. NO. 011-42-3915 | MAR. 12, 1973 NAME JACOBS ANTHONY WILLIAMS | BIRTH DATE 03-27-42 | COL 02

1 5 | STREET ADDRESS 910 N FRANCISCO | CITY CHICAGO | ST. IL | ZIP 60604

H S CODE | PARENTS NAME JACOBS FRANK A | TELEPHONE 312 479-3512

CLASS RANK | HIGH SCHOOL STEVENS INSTITUTE | YR. 60 | ENTER 2 70

CLASS SIZE | DEC | SAT V | SAT M | DEPAUL V | DEPAUL Q | CLS | FEE | ADMITTED | LETTER | CREDENTIALS

PREVIOUS DEGREE | DEGREE | MAJOR | HON | YEAR

COLLEGE OR UNIVERSITY

SOC. SEC. NO. 061-49-3861 | MAR. 12, 1973 NAME CAROLLA ANN COURTNEY | BIRTH DATE 03-21-35 | COL 15

1 5 | STREET ADDRESS 2514 BELLAMY BLVD | CITY EVANSTON | ST. IL | ZIP 60201

H S CODE | PARENTS NAME SISTERS OF PROVIDENCE | TELEPHONE 312 876-7370

CLASS RANK | HIGH SCHOOL CHEVERUS MALDEN MA | YR. 53 | ENTER 4 72

CLASS SIZE | DEC | SAT V | SAT M | DEPAUL V | DEPAUL Q | CLS | FEE | ADMITTED | LETTER | CREDENTIALS

PREVIOUS DEGREE | DEGREE | MAJOR | HON | YEAR

COLLEGE OR UNIVERSITY

FIGURE 7

* Param 4 - Table cards for Data FieTd Expansion/Conversion feature.
* Param 5 - Record Selection information. Although not shown in example, this capability is similar to General Extract feature.
* Param 6 - Control Break, Summarizing and Crossfooting specifications.

General Print matches the report records on the input tape to the parameter cards by report code (AA on beginning of tape, for example). When a Param 1 card report code is found on the tape, the rest of the parameter cards for the report are read in and analyzed (stops reading cards when next Param 1 card is read) and report printing begins.

At the end of a report, the Param 1 card last entered is matched to the next report record and the process continues until the end of tape is read where run statistics are printed.

It is through pre-programmed routines like General Extract and Print that the DePaul System can accommodate widely varied needs and can quickly respond to changes as they occur.